

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-255881

(43)Date of publication of application : 21.09.2001

(51)Int.Cl. G10L 13/00
G06F 3/16
G06F 13/00
G10L 15/00
G10L 15/28
G10L 15/22

(21)Application number : 2000-068992

(71)Applicant : MATSUSHITA ELECTRIC IND CO LTD

(22)Date of filing : 13.03.2000

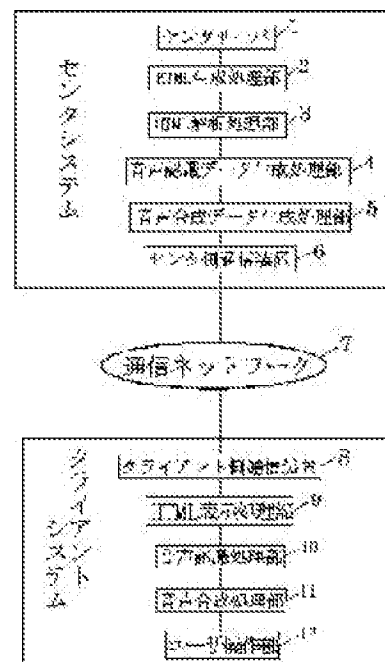
(72)Inventor : TAKAYANAGI YUICHI
UNO KASHU

(54) AUTOMATIC SPEECH RECOGNITION/SYNTHESIS BROWSER SYSTEM

(57)Abstract:

PROBLEM TO BE SOLVED: To conduct speech operations by means of speech recognition/speech synthesis without changing a client system in accordance with the service of a center side.

SOLUTION: An HTML sentence which matches with service providing contents is generated by an HTML generation processing section 2 by the instruction of a center server 1. Data analysis of the HTML is conducted by an HTML analysis processing section 3. Speech recognition data and speech synthesis data matched with the service are generated by a speech recognition data generating section 4 and a speech synthesis data generating section 5. Interpretation of the received HTML data is conducted by an HTML display processing section 8 of a client system such as an automobile navigation device, a personal computer and a PDA. Speech inputting and outputting are conducted in a user operation section 12 employing the speech recognition and the speech synthesis data so that user's speech operations are made possible. Thus, the service of the center is expanded without changing the client's system.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-255881

(P2001-255881A)

(43) 公開日 平成13年9月21日 (2001.9.21)

(51) Int.Cl. ⁷	識別記号	F I	テ-マコード* (参考)
G 1 0 L 13/00		G 0 6 F 3/16	3 4 0 A 5 D 0 1 5
G 0 6 F 3/16	3 4 0		3 4 0 N 5 D 0 4 5
		13/00	5 5 0 A 9 A 0 0 1
13/00	5 5 0	G 1 0 L 3/00	R
G 1 0 L 15/00			E

審査請求 未請求 請求項の数 6 O L (全 8 頁) 最終頁に続く

(21) 出願番号 特願2000-68992(P2000-68992)

(22) 出願日 平成12年3月13日 (2000.3.13)

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 ▲高▼▲柳▼ 雄一

神奈川県横浜市港北区綱島東四丁目3番1

号 松下通信工業株式会社内

(72) 発明者 宇野 嘉修

神奈川県横浜市港北区綱島東四丁目3番1

号 松下通信工業株式会社内

(74) 代理人 100099254

弁理士 役 昌明 (外3名)

Fターム(参考) 5D015 AA04 BB01 KK01 KK04 LL06

5D045 AA01 AB01 AB17 AB21 AB24

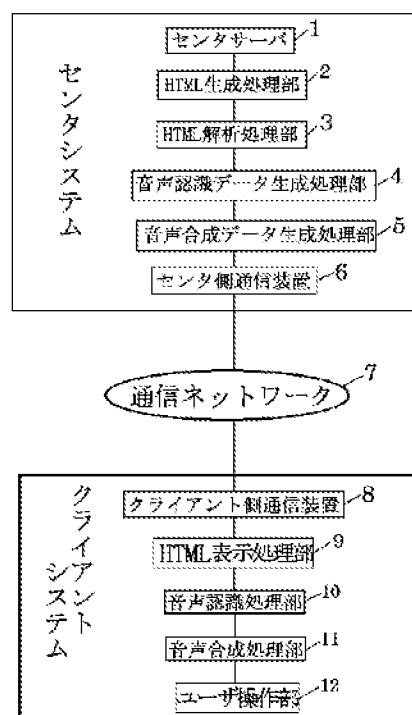
9A001 HH17 HH18 JJ18 JJ25 JJ27

(54) 【発明の名称】 自動音声認識／合成ブラウザシステム

(57) 【要約】

【課題】 センタ側のサービスに合わせてクライアントシステムを変えることなく、音声認識・音声合成による音声操作を可能とする。

【解決手段】 センタサーバ1の指示で、サービス提供内容に合わせたHTML文を、HTML生成処理部2で生成する。HTML解析処理部3で、HTMLのデータ解析を行ない、音声認識データ生成部4と音声合成データ生成部5で、サービスに合わせた音声認識データと音声合成データを生成する。カーナビゲーション装置やパソコンやPDAなどのクライアントシステムのHTML表示処理部8で、受信したHTMLデータの解釈を行なう。ユーザ操作部12で、音声認識データと音声合成データを使って音声入出力を行ない、ユーザの音声操作を可能とする。クライアントシステムを変えることなく、センタのサービスの拡張ができる。



【特許請求の範囲】

【請求項1】 センタ装置とクライアント装置と通信ネットワークとからなる自動音声認識／合成ブラウザシステムにおいて、前記センタ装置に、センタが提供するサービスを実行するサーバ装置と、前記サービスに対応したHTML文を生成するHTML文生成手段と、前記HTML文を解析するHTML文解析手段と、前記HTML文の解析結果に応じて音声認識データを生成する音声認識データ生成手段と、前記HTML文の解析結果に応じて音声合成データを生成する音声合成データ生成手段と、前記HTML文と前記音声認識データと前記音声合成データとを前記通信ネットワークを介して前記クライアント装置に送信するセンタ側通信手段とを設け、前記クライアント装置に、前記HTML文と前記音声認識データと前記音声合成データとを前記センタ装置から前記通信ネットワークを介して受信するクライアント側通信手段と、前記HTML文を処理して表示するHTML文表示手段と、前記HTML文と前記音声認識データとに基づいて音声認識処理を行なう音声認識手段と、前記HTML文と前記音声合成データとに基づいて音声合成処理を行なう音声合成手段と、前記音声認識処理の結果に基づいて入力操作を行なうユーザ操作手段とを設けたことを特徴とする自動音声認識／合成ブラウザシステム。

【請求項2】 前記HTML文生成手段に、前記HTML文の中で新規なデータ構造を定義する手段を設けたことを特徴とする請求項1記載の自動音声認識／合成ブラウザシステム。

【請求項3】 前記音声認識データ生成手段に、前記音声認識データを動的に変更する手段を設け、前記音声合成データ生成手段に、前記音声合成データを動的に変更する手段を設けたことを特徴とする請求項1記載の自動音声認識／合成ブラウザシステム。

【請求項4】 前記HTML文生成手段に、HTMLの標準タグのみを用いてHTML文を生成する手段を設けたことを特徴とする請求項1記載の自動音声認識／合成ブラウザシステム。

【請求項5】 前記ユーザ操作手段に、前記音声認識の結果に基づいてキー押下操作を実行する手段を設けたことを特徴とする請求項1記載の自動音声認識／合成ブラウザシステム。

【請求項6】 前記センタ側通信手段に、ADPCMファイルを送信する手段を設け、前記クライアント側通信手段に、前記ADPCMファイルを受信する手段を設け、前記音声合成手段に、前記ADPCMファイルを再生する手段を設けたことを特徴とする請求項1記載の自動音声認識／合成ブラウザシステム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、自動音声認識／合成ブラウザシステムに関し、特に、ブラウザ搭載シス

テムにおいて任意の用語を音声認識・合成できる自動音声認識／合成ブラウザシステムに関する。

【0002】

【従来の技術】 従来のブラウザ搭載システムでは、音声認識合成辞書や語彙集を格納しておき、それらの辞書や語彙集に合わせて、音声認識合成を行なっている。新たな辞書の作成時や語彙集の作成時には、ブラウザ搭載システム内の音声合成用ROMや辞書格納補助記憶装置のメディアを更新している。

【0003】 このようなシステムの従来例として、特開平11-249867号公報に開示されている「音声ブラウザシステム」がある。このシステムは、図3に示すように、視覚障害者であってもWWW情報を取得することが可能なシステムである。このシステムでは、ユーザの音声による要求を入力すると、サーバで音声入力による要求を音声認識する。クライアントは、サーバから取得したURLに基づいて、サーバにHTMLファイル要求する。サーバは、クライアントから指定されたURLに対して、要求をインターネットに送信する。インターネットから取得した応答から音読テキストを抽出し、音声データに変換して音声合成する。生成された音声データをクライアントに送信し、クライアント側で音声データを出力する。

【0004】 また、インターネットが普及している今日では、様々なサービスが、センタ側で整備されてきている。例えば、タクシーサービスにおける音声認識データ／音声合成データ群がある。水道やマンホールの整備をするサービスにおける音声認識データ群及び音声データ群がある。

【0005】

【発明が解決しようとする課題】 しかし、上記従来のブラウザ搭載システムでは、センタ側のサービスの多様化に合わせて、クライアント側の音声認識や音声合成のシステムを変更しなければならないという問題がある。

【0006】 本発明は、上記従来の問題を解決して、センタ側のサービスに合わせてクライアントシステムを変えることなく、音声認識・音声合成による音声操作を可能とする自動音声認識／合成ブラウザシステムを提供することを目的とする。

【0007】

【課題を解決するための手段】 上記の課題を解決するために、本発明では、センタ装置とクライアント装置と通信ネットワークとからなる自動音声認識／合成ブラウザシステムのセンタ装置に、センタが提供するサービスを実行するサーバ装置と、サービスに対応したHTML文を生成するHTML文生成手段と、HTML文を解析するHTML文解析手段と、HTML文の解析結果に応じて音声認識データを生成する音声認識データ生成手段と、HTML文の解析結果に応じて音声合成データを生成する音声合成データ生成手段と、HTML文と音声認

識データと音声合成データとを通信ネットワークを介してクライアント装置に送信するセンタ側通信手段とを設け、クライアント装置に、HTML文と音声認識データと音声合成データとをセンタ装置から通信ネットワークを介して受信するクライアント側通信手段と、HTML文を処理して表示するHTML文表示手段と、HTML文と音声認識データとに基づいて音声認識処理を行なう音声認識手段と、HTML文と音声合成データとに基づいて音声合成処理を行なう音声合成手段と、音声認識処理の結果に基づいて入力操作を行なうユーザ操作手段とを設けた構成とした。このように構成したことにより、センタのサービスの種別を利用者が意識することなく、センタから送られた画面上で音声操作ができる。

【0008】また、音声認識データ生成手段に、音声認識データを動的に変更する手段を設け、音声合成データ生成手段に、音声合成データを動的に変更する手段を設けたので、クライアント側のパソコンやカーナビゲーションの音声辞書・音声合成語彙ファイルを修正する必要がなく、全てのサービスに対して同じクライアントシステムを用いることができる。

【0009】

【発明の実施の形態】以下、本発明の実施の形態について、図1と図2を参照しながら詳細に説明する。

【0010】（実施の形態）本発明の実施の形態は、センタ装置で、サービスに合わせたHTML文と音声認識データと音声合成データを生成し、カーナビゲーション装置やパソコンやPDAなどのクライアント装置で、受信したHTML文と音声認識データと音声合成データを使って音声入出力を行ない、ユーザの音声操作を可能とする自動音声認識／合成ブラウザシステムである。

【0011】図1は、本発明の実施の形態における自動音声認識／合成ブラウザシステムの機能ブロック図である。図1において、センタサーバ1は、駐車場管理センタやドライブスルー店舗などにおける様々なサービスを行なう情報処理装置である。HTML生成処理部2は、HTMLのデータ生成を行なう処理部である。HTML解析処理部3は、HTMLのデータ解析を行なう処理部である。音声認識データ生成処理部4は、音声認識用のデータ生成を行なう処理部である。音声合成データ生成処理部5は、音声合成用のデータ生成を行なう処理部である。センタ側通信装置6は、センタと通信ネットワークを結ぶ通信手段である。通信ネットワーク7は、公衆電話回線やインターネットなどの通信回線である。クライアント側通信装置8は、パソコンやカーナビゲーションやPDAなどのクライアントシステムでHTMLデータを受信する通信手段である。HTML表示処理部9は、HTML文を解釈して表示する処理部である。音声認識処理部10は、ユーザの音声を認識する手段である。音声合成処理部11は、HTML文の音声合成データを音声にして出力する手段である。ユーザ操作部12は、ユー

ザが音声で装置の操作を行なう手段である。

【0012】図2は、自動音声認識／合成ブラウザの画面イメージを示す図である。

【0013】上記のように構成された本発明の実施の形態における自動音声認識／合成ブラウザシステムの動作を説明する。図1に示すように、センタサーバ1が、サービス提供内容に合わせたHTML文を生成するように、HTML生成処理部2に指示する。HTML生成処理部2は、その指示に基づいてHTML文を生成する。HTML解析処理部3は、生成されたHTML文のデータ解析を行なう。音声認識データ生成処理部4は、HTML文の解析結果に基づいて、音声認識用のデータ生成を行なう。音声合成データ生成処理部5は、HTML文の解析結果に基づいて、音声合成用のデータ生成を行なう。センタ側通信装置6は、通信ネットワーク7を介してHTML文と音声認識データと音声合成データをクライアントシステムに送信する。

【0014】車載ナビゲーションシステムやパソコンなどのクライアントシステムの通信装置8は、センタからのHTML文と音声認識データと音声合成データを受信する。HTML表示処理部9は、受信したHTML文を解釈して表示を行なう。音声認識処理部10は、受信したHTML文と音声認識データに基づき、ユーザの音声を認識する。音声合成処理部11は、受信したHTML文と音声合成データに基づき、音声合成を行なう。ユーザ操作部8は、音声認識結果に基づいて、操作入力処理を行なう。

【0015】図2を参照して、ハンバーガーショップでの注文システムを例にして、動作の流れを説明する。この画面が表示されたとき、ユーザは、「ハンバーガー／セットメニュー」か、「サイドメニュー」か、「ドリンク／デザート」のいずれかを発声する。音声認識されると、それに対応したページにジャンプする。「サイドメニュー」選択時には、そのページに移行し、「ドリンク／デザート」選択時は、そのページに移行する。この画面において、「ハンバーガーアイコン」と発声することにより、ハンバーガーの個数のエリアに1が入力される。「オーダーの確認」と発声することにより、オーダーが終了する。

【0016】音声認識／合成情報の記述例を示すと、次のようになる。

```
<OBJECT classid="VOICERECOG" width="1" height="1">
<PARAM name="guide" value="TEXT:当店へようこそ。メニューから希望の商品と数量を選択してください。ADPCM:MAC_WELCOM,ADPCM:MAC_ORDER">
<PARAM name="recog" value="KEY:0,REC:ハンバーガーセットメニュー,REC:セットメニュー,TEXT:ハンバーガー／セットメニュー">
<PARAM name="recog" value="KEY:1,REC:サイドメニュー,TEXT:サイドメニュー">
```

<PARAM name="recog" value="KEY:2,REC:ドリンクデザートメニュー,REC:ドリンクメニュー,REC:デザートメニュー,TEXT:ドリンク/デザートメニュー">

</OBJECT>

音声認識(VOICERECOG)情報の定義方法を説明する。フォーマットは、

<OBJECT classid="VOICERECOG" width="幅" height="高さ">

ようになる。width="幅"で、オブジェクトの幅を指定する。"幅"は、常に"1"を指定する。省略はできない。height="高さ"で、オブジェクトの高さを指定する。"高さ"は、常に"1"を指定する。省略はできない。幅と高さに"1"以上の値を指定することも可能であるが、無駄な空白がレイアウトされてしまうので、必ず"1"を指定する。幅と高さに"0"を指定すると、音声認識モジュールは起動されない。次の例

<OBJECT classid="VOICERECOG" width="1" height="1">

では、1×1の表示サイズを持つ音声認識情報が定義される。

【0017】音声認識モジュールが起動されるのは、音声認識情報の幅と高さで指定した領域が、画面に表示されているときだけである。同一ページ内に音声認識情報を複数定義しても、それぞれが同時に画面内に表示されなければ、音声認識モジュールは正常に動作する。同時に複数の音声認識情報が画面に表示されるようにページを記述すると、最後に表示された音声認識情報を使用して音声認識モジュールが動作する。

【0018】音声認識(guide)説明文の追加方法を説明する。フォーマットは、

<PARAM name="guide" value="TEXT:表示文字,ADPCM:ADPCMファイル名,VOICE:テキスト合成音声,MONEY:金額">

ようになる。【0019】value="説明文の定義"で、説明文を定義する。valueの各エントリの区切りには、","文字を使用する。","文字から次のエントリの開始まで、半角スペース、タブ、改行コードを任意に挿入できる。TEXT:表示文字で、説明文の表示文字を定義する。表示文字の定義を省略した場合は、デフォルトの表示文字列として、

"音声認識を開始します。"が使用される。複数表示文字が定義された場合は、後に定義されたものが有効となる。ADPCM:ADPCMファイル名で、ADPCM音声をファイル名で定義する。VOICE:テキスト合成音声で、テキスト合成音声を全角カナ文字で定義する。MONEY:金額で、"0"~"9"の数字を使用して、金額を定義する。例えば、"400"は、「ヨンヒャク」と発声される。"guide"エントリは省略可能である。省略した場合は、次に示すデフォルトの音声認識処理の説明文

<PARAM name="guide" value="TEXT:音声認識を開始します。VOICE:オンセイニンシキヲカシシマス">

が使用される。"guide"エントリが複数定義された場合は、後に定義されたものが有効となる。ADPCM:とVOICE:は複数定義でき、定義順に再生される。"guide"エントリ内に最低一つは、"ADPCM:ADPCMファイル名"、もしくは"VOICE:テキスト合成音声"を定義する必要がある。

【0020】次の例

<PARAM name="guide" value="TEXT:当店へようこそ。メニューから希望の商品と数量を選択してください。ADPCM:MAC_WELCOM,ADPCM:MAC_ORDER">

では、表示文字列として、「当店へようこそ。メニューから希望の商品と数量を選択してください。」が表示され、ADPCMファイルが、"MAC_WELCOM"、"MAC_ORDER"の順で再生される。

【0021】音声認識(recog)単語を登録する方法を説明する。フォーマットは、

<PARAM name="recog" value="KEY:アクセスキー,REC:認識文字,TEXT:表示文字">

ようになる。

【0022】value="認識単語の定義"で、認識単語を定義する。valueの各エントリの区切りには、","文字を使用する。","文字から次のエントリの開始まで、半角スペース、タブ、改行コードを任意に挿入できる。KEY:アクセスキーで、認識単語が一致したときに反応するアクセスキーを定義する。省略不可である。REC:認識文字で、認識文字を全角カナ文字で定義する。"recog"エントリ内に複数定義可能である。省略不可である。TEXT:表示文字で、表示文字を定義する。省略した場合は、認識単語の説明文は表示されない。複数定義した場合は、後に定義された表示文字が使用される。「ハンバーガーイッコ」~「ハンバーガージュッコ」など、あまりにも認識単語が増える場合は、例1~3の様に、「ハンバーガー」という基本単語の表示文字だけを定義すればよい。"recog"エントリは、複数定義可能である。実装系によって、"recog"エントリの最大数は異なる。同一の認識文字を複数のアクセスキーに対応付けする様な定義を行った場合、動作は保証されない。

【0023】例1

<PARAM name="recog" value="KEY:2,REC:ハンバーガー,TEXT:ハンバーガー">

では、説明文として、「ハンバーガー」が表示され、「ハンバーガー」が認識されると、アクセスキー"2"に対応したアイテムが反応する。

【0024】例2

<PARAM name="recog" value="KEY:101,REC:ハンバーガーイッコ">

<PARAM name="recog" value="KEY:101,REC:ハンバーガーヒトツ">

では、説明文は表示されず、「ハンバーガーイッコ」、「ハンバーガーヒトツ」が認識されると、アクセスキー"101"に対応したアイテムが反応する。

【0025】例3

<PARAM name="recog" value="KEY:101,REC:ハンバーガー
イッコ,REC:ハンバーガーヒトツ">

では、例2と同様の動作を行う。HTMLのサイズを小さくしたい時は、このように記述するとよい。

【0026】音声合成(VOICECOMPOSE)情報の定義方法を説明する。フォーマットは、

<OBJECT classid="VOICECOMPOSE" width="幅" height="高さ">

のようになる。

【0027】width="幅"で、オブジェクトの幅を指定する。"1"を指定する。省略不可である。height="高さ"で、オブジェクトの高さを指定する。"1"を指定する。省略不可である。幅と高さに"1"以上の値を指定することも可能であるが、無駄な空白がレイアウトされてしまうだけである。必ず"1"を指定する。幅と高さに"0"を指定すると、音声合成モジュールは起動されない。音声合成モジュールが起動されるのは、音声合成情報の幅と高さで指定した領域が画面に表示されているときだけである。同一ページ内に音声合成情報を複数定義しても、それぞれが同時に画面内に表示されなければ、音声合成モジュールは正常に動作する。

【0028】次の例

<OBJECT classid="VOICECOMPOSE" width="1" height="1">

では、1×1の表示サイズを持つ音声合成情報が定義される。

【0029】音声合成(guide)の定義方法を説明する。

フォーマットは、

<PARAM name="guide" value="TEXT:表示文字,ADPCM:ADPCMファイル名,VOICE:テキスト合成音声,MONEY:金額">

のようになる。

【0030】value="合成音声の定義"で、合成音声を定義する。valueの各エントリの区切りには、","文字を使用する。","文字から次のエントリの開始まで、半角スペース、タブ、改行コードを任意に挿入できる。TEXT:表示文字で、合成音声の表示文字を定義する。"TEXT:表示文字"は、省略可能である。省略した場合、合成音声の説明文は表示されない。複数の"TEXT:表示文字"が定義された場合は、後に定義されたものが有効となる。ADPCM:ADPCMファイル名で、ADPCM音声をファイル名で定義する。VOICE:テキスト合成音声で、テキスト合成音声を全角カナ文字で定義する。MONEY:金額で、"0"~"9"の数字を使用して、金額を定義する。例えば、"400"は、「ヨンヒャク」と発声される。

【0031】"guide"エントリは、省略不可である。"guide"エントリが複数定義された場合は、後に定義されたものが有効となる。ADPCM:とVOICE:は、複数定義でき、定義順に再生される。"guide"エントリ内に最低一つは、"ADPCM:ADPCMファイル名"、もしくは"VOICE:テキス

ト合成音声"を定義する必要がある。

【0032】次の例

<PARAM name="guide" value="TEXT:当店へようこそ。メニューから希望の商品と数量を選択してください。ADPCM:MAC_WELCOM,ADPCM:MAC_ORDER">

では、表示文字列として、「当店へようこそ。メニューから希望の商品と数量を選択してください。」が表示され、ADPCMファイルが、"MAC_WELCOM"、"MAC_ORDER"の順で再生される。

10 【0033】"アクセスキー"の通知によって反応するアイテムの定義方法を説明する。フォーマットは、accesskey="アクセスキー"である。

【0034】"アクセスキー"には、"a"~"z"、"A"~"Z"、"0"~"9"の英数字キーに加え、"10"~"255"の数字を割り当てることができる。アクセスキーは、画面に表示されたアイテムにのみ通知される。アクセスキーが定義されていても、画面に表示されていないければ、アクセスキーの通知に反応できない。

【0035】次の例

20
ハンバーガー/セットメニュー

では、アクセスキー"0"が反応すると、"#burger1"にページ内ジャンプする。

【0036】次の例

<SELECT name="quantity_burger1" size="1" accesskey="2">

<OPTION accesskey="100" value="0">0

<OPTION accesskey="101" value="1">1

<OPTION accesskey="102" value="2">2

30 <OPTION accesskey="103" value="3">3

<OPTION accesskey="104" value="4">4

<OPTION accesskey="105" value="5">5

<OPTION accesskey="106" value="6">6

<OPTION accesskey="107" value="7">7

<OPTION accesskey="108" value="8">8

<OPTION accesskey="109" value="9">9

<OPTION accesskey="110" value="10">10

</SELECT>

では、アクセスキー"2"が反応すると、セレクトアイテム"quantity_burger1"がプルダウン表示される。アクセスキー"100"~"110"が反応すると、0~10のセレクトアイテムがプルダウン表示されず選択される。

【0037】このようにすることにより、クライアント側の音声認識合成システムを変える必要がなく、センタ側のコンテンツを変更することにより、サービスに応じたユーザの音声操作が可能となる。カタカナ文字からの音声合成ならびにADPCMファイルの再生も可能である。インターネットの標準言語であるHTMLを有効に活用して、サービスごとの開発費用の削減と、開発効率の向上が実現できる。サービスの拡張が容易にでき、サ

ービスの多様化に対応できる。

【0038】上記のように、本発明の実施の形態では、自動音声認識／合成ブラウザシステムを、センタシステムで、サービスに合わせたHTML文と音声認識データと音声合成データを生成し、カーナビゲーション装置やパソコンやPDAなどのクライアントシステムで、受信したHTML文と音声認識データと音声合成データを使って音声入出力を行ない、ユーザの音声操作を可能とする構成としたので、クライアント側のシステムを変えることなく、ユーザの音声操作を可能とし、センタ側のコンテンツを可変にすることにより、サービスの多様化に対応できる。

【0039】

【発明の効果】以上の説明から明らかなように、本発明の自動音声認識／合成ブラウザシステムでは、センタ装置とクライアント装置と通信ネットワークとからなる自動音声認識／合成ブラウザシステムのセンタ装置に、センタが提供するサービスを実行するサーバ装置と、サービスに対応したHTML文を生成するHTML文生成手段と、HTML文を解析するHTML文解析手段と、HTML文の解析結果に応じて音声認識データを生成する音声認識データ生成手段と、HTML文の解析結果に応じて音声合成データを生成する音声合成データ生成手段と、HTML文と音声認識データと音声合成データとを通信ネットワークを介してクライアント装置に送信するセンタ側通信手段とを設け、クライアント装置に、HTML文と音声認識データと音声合成データとをセンタ装置から通信ネットワークを介して受信するクライアント*

*側通信手段と、HTML文を処理して表示するHTML文表示手段と、HTML文と音声認識データとに基づいて音声認識処理を行なう音声認識手段と、HTML文と音声合成データとに基づいて音声合成処理を行なう音声合成手段と、音声認識処理の結果に基づいて入力操作を行なうユーザ操作手段とを設けた構成としたので、センタのサービス種別を利用者が意識することなく、サービスに対応した音声操作が極めて簡単にできるという効果が得られる。

10 【図面の簡単な説明】

【図1】本発明の実施の形態における自動音声認識／合成ブラウザシステムの機能ブロック図、

【図2】本発明の実施の形態における画面例、

【図3】従来の音声／合成ブラウザシステムの機能ブロック図である。

【符号の説明】

- 1 センタサーバ
- 2 HTML生成処理部
- 3 HTML解析処理部
- 4 音声認識データ生成処理部
- 5 音声合成データ生成処理部
- 6 センタ側通信装置
- 7 通信ネットワーク
- 8 クライアント側通信装置
- 9 HTML表示処理部
- 10 音声認識処理部
- 11 音声合成処理部
- 12 ユーザ操作部

【図2】

ハンバーガー／セットメニュー

サイドメニュー

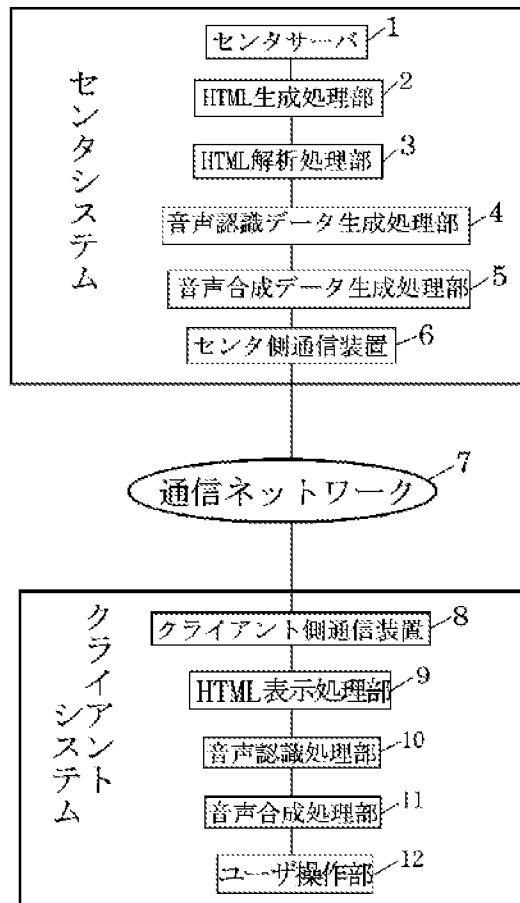
ドリンク／デザート

- ハンバーガー ¥100 × 個
- チーズバーガー ¥130 × 個
- テリヤキバーガー ¥190 × 個
- チキンバーガー ¥190 × 個

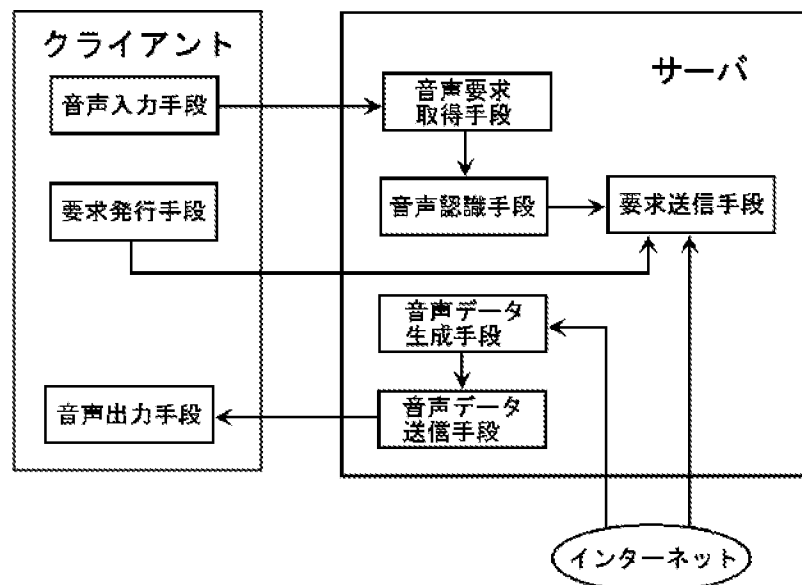
▼ ハンバーガー／セットメニューへ戻る ▼

オーダー内容の確認 / 完了

【図1】



【図3】



フロントページの続き

(51)Int.Cl. ⁷	識別記号	F I	ターコード (参考)
G 1 0 L 15/28		G 1 0 L 3/00	5 5 1 A
15/22			5 5 1 P
			5 6 1 C